Peeter Joot
peeter.joot@gmail.com

ECE1254H Modeling of Multiphysics Systems. Lecture 5: Numerical error and conditioning. Taught by Prof. Piero Triverio

1.1 Numerical errors and conditioning

1.1.1 *Strict diagonal dominance*

Related to a theorem on one of the slides:

Definition 1.1: Strictly diagonally dominant

A matrix $[M_{ij}]$ is strictly diagonally dominant if

$$|M_{ii}| > \sum_{j \ne i} |M_{ij}| \quad \forall i \tag{1.1}$$

For example, the stamp matrix

$$\begin{matrix} & i & j \\ \begin{matrix} i \\ j \end{matrix} & \multicolumn{2}{c}{\begin{bmatrix} \frac{1}{R} & -\frac{1}{R} \\ -\frac{1}{R} & \frac{1}{R} \end{bmatrix}} \end{matrix} \tag{1.2}$$

is not strictly diagonally dominant. For row i this strict dominance can be achieved by adding a reference resistor

$$\begin{matrix} & i & j \\ \begin{matrix} i \\ j \end{matrix} & \multicolumn{2}{c}{\begin{bmatrix} \frac{1}{R_0} + \frac{1}{R} & -\frac{1}{R} \\ -\frac{1}{R} & \frac{1}{R} \end{bmatrix}} \end{matrix} \tag{1.3}$$

However, even with strict dominance, we can have trouble with ill posed (perturbative) systems. Round off error examples with double precision

$$(1 - 1) + \pi 10^{-17} = \pi 10^{-17}, \tag{1.4}$$

vs.

$$\left(1 + \pi 10^{-17}\right) - 1 = 0. \tag{1.5}$$

This is demonstrated by

```
#include <stdio.h>
#include <math.h>

// produces:
// 0 3.14159e-17
int main()
{
   double d1 = (1 + M_PI * 1e-17) - 1 ;
   double d2 = M_PI * 1e-17 ;

   printf( "%g %g\n", d1, d2 ) ;

   return 0 ;
}
```

Note that a union and bitfield [1] can be useful for exploring double precision representation.

1.1.2 *Exploring uniqueness and existence*

For a matrix system $\overline{M}x = \bar{b}$ in column format, with

$$\begin{bmatrix} \overline{M}_1 & \overline{M}_2 & \cdots & \overline{M}_N \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \bar{b}. \tag{1.6}$$

This can be written as

$$\underbrace{\boxed{x_1}}_{\text{weight}} \overline{M}_1 + \underbrace{\boxed{x_2}}_{\text{weight}} \overline{M}_2 + \cdots x_N \overline{M}_N = \bar{b}. \tag{1.7}$$

Linear dependence means

$$y_1 \overline{M}_1 + y_2 \overline{M}_2 + \cdots y_N \overline{M}_N = 0, \tag{1.8}$$

or $M\bar{y} = 0$.

With a linear dependency an additional solution, given solution $\bar{x}$ is $\bar{x}^1 = \bar{x} + \alpha y$. This becomes relevant for numerical processing since for a system $M\bar{x}^1 = \bar{b}$ we can often find $\alpha M\bar{y}$, for which

$$M\bar{x} + \alpha M\bar{y} = \bar{b}, \tag{1.9}$$

where $\alpha M\bar{y}$ is of order 10^{-20}.

Table 1.1: Solution space

	$\bar{b} \in \text{span}\{M\_i\}$	$\bar{b} \notin \text{span}\{M\_i\}$
columns of M linearly independent	$\bar{x}$ exists and is unique	No solution
columns of M linearly dependent	$\bar{x}$ exists. Infinitely many solutions	No solution

1.1.3 *Perturbation and norms*

Consider a perturbation to the system $M\bar{x} = \bar{b}$

$$(M + \delta M)(\bar{x} + \delta \bar{x}) = \bar{b}. \tag{1.10}$$

Some vector norms

- L_1 norm

$$\|\bar{x}\|_1 = \sum_i |x_i| \tag{1.11}$$

- L_2 norm

$$\|\bar{x}\|_2 = \sqrt{\sum_i |x_i|^2} \tag{1.12}$$

- L_∞ norm

$$\|\bar{x}\|_\infty = \max_i |x_i|. \tag{1.13}$$

These are illustrated for $\bar{x} = (x_1, x_2)$ in fig. 1.1.

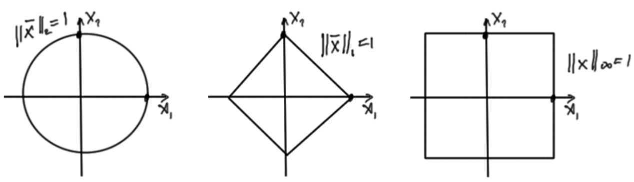

Figure 1.1: Some vector norms

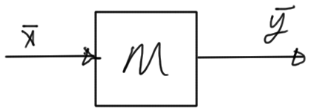

Figure 1.2: Matrix as a transformation

1.1.4 *Matrix norm*

A matrix operation $\bar{y} = M\bar{x}$ can be thought of as a transformation as in fig. 1.2.

The 1-norm for a Matrix is defined as

$$\|M\| = \max_{\|\bar{x}\|_1=1} \|M\bar{x}\|\,, \tag{1.14}$$

and the matrix 2-norm is defined as

$$\|M\|_2 = \max_{\|\bar{x}\|_2=1} \|M\bar{x}\|_2\,. \tag{1.15}$$

Bibliography

[1] Peeter Joot. *Simple C++ double representation explorer*, 2014. URL https://github.com/peeterjoot/physicsplay/blob/master/notes/ece1254/samples/rounding.cpp. [Online; accessed 29-Sept-2014]. 1.1.1